I

Introduction and summary

1 THE NARRATIVE IN A NUTSHELL AND THE MORAL OF THE TALE

Nationalism demands that rulers and ruled hail from the same ethnic background. The gradual adoption of this principle of legitimate statehood has transformed the shape of the political world over the past 200 years and has provided the ideological motivation for an increasing number of wars fought in the modern era. Before the age of nationalism set in at the end of the eighteenth century, individuals did not pay much attention to their own ethnic background or that of their rulers. They identified primarily with a local community – a village or town, a clan, or a mosque. In much of Europe and East Asia, their overlords ruled in the name of a divine dynasty, rather than "the people," and many were of different ethnic stock than their subjects. In parts of the Middle East, Africa, or Central Asia, charismatic leaders held tribal confederacies together and were respected and feared for their political skills and military bravery. Vast stretches of land in the Americas, the Middle East, and Eastern Europe were ruled by emperors whose legitimacy derived from spreading God's word across the world (as did the Ottomans and Bourbons) or bringing civilization to "backward" peoples (as France and Great Britain claimed to do in their colonies). At the beginning of the nineteenth century, such empires covered about half of the world's surface, while dynastic kingdoms, tribal confederacies, city-states, and so forth, made up most of the rest, as Figure 1.1 shows.

    In this world of empires, dynastic kingdoms, city-states, and tribal confederacies, few wars concerned the ethno-national composition of government. Rather, they were fought by dynastic states over the balance of power between them or over the rightful successor to a throne. Empires conquered fertile lands
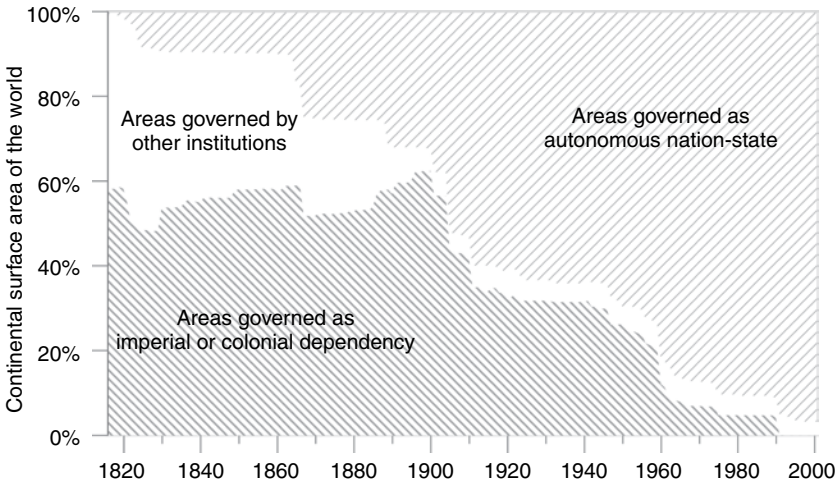
FIGURE 1.1 Empires, nation-states, and other types of polities, 1816–2001
Notes: states smaller than 25,000 km² are excluded; data are from Chapter 4.

far away from their capitals. Alliances of city-states competed over trade routes or rural hinterlands. Rebellious movements saw to bring heavenly order to the corrupt politics of the day or to repeal an unjust tax increase. At the beginning of the nineteenth century, still only one-fourth of the wars were ethno-nationalist, as can be seen from Figure 1.2, while balance-of-power wars between states, wars of conquest, and non-ethnic civil wars each comprised another quarter of all violent conflicts.

A contemporary observer looks at a different world and through different eyes. The globe is divided into a series of sovereign states, each supposed to represent a nation bound together by shared history and common culture. To us, this political map seems as obvious as the shapes of continents and the rivers that run through them. With the exception of the Middle Eastern monarchies and some small European principalities, most of today's states are ruled in the name of a nation of equal citizens, rather than dynasty or divine will. Statehood has become so much associated with nationalist principles that the terms nations and states are often used interchangeably, as in the "United Nations" or in "inter-national."

Most of today's more prominent and protracted wars are also associated with the national principle – the idea that each people should be self-ruled, that ethnic like should be governed by like. The independence struggle of Abkhazians against the Georgian state or the conflict between Protestant and Catholic parties and militias in Northern Ireland come to mind. Figure 1.2 shows that at the end of the twentieth century, over three-quarters of all full-scale wars – those armed conflicts costing more than 1,000 battle deaths – were fought either by
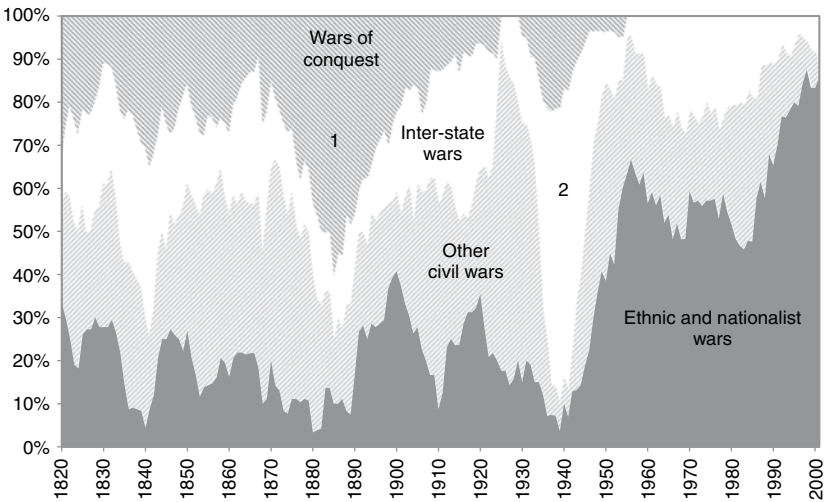
FIGURE I.2   The ethno-nationalization of war, 1816–2001
Notes: ten-year moving averages; for data sources see Chapter 4; 1 marks wars associated with the conquests of Africa and Central Asia; 2 marks wars associated with World War II.

nationalists who seek to establish a separate nation-state or over the ethnic balance of power within an existing state. Contrary to what Karl Marx had predicted, the twentieth century has turned into the age of ethno-nationalist conflict, rather than revolutionary class struggle.

This book seeks to explain this momentous transformation of the political world – from a world of multiethnic empires, dynastic kingdoms, tribal confederacies, and city-states to a world of states each ruled in the name of a nation properly seated in the general assembly of the UN; from localized political identities to large-scale ethnic or national communities with often millions, sometimes tens of millions, of members; from wars of conquest, succession, and tax rebellions to wars in the name of national sovereignty and grandeur, ethnic autonomy, and the like.

How has this transformation come about? Existing scholarship has mostly focused on how strong, territorially centralized states have emerged in Western Europe and beyond. Charles Tilly's famed dictum that "wars made states and states made war" referred to the rise of these absolutist states from the sixteenth to the eighteenth centuries. This book takes this story from the early modern period into our present day and from Western Europe to the world. It is not concerned with the development of the sovereign territorial state, as were Tilly and his successors, but why these states became nation-states and how this particular model of legitimizing political power proliferated across the world. It shows that

the shift from dynasticism and empire to the nation-state was both the cause and consequence of a new wave of wars long after early modern states had been formed in previous centuries of warfare. This new wave, carried forward by the power of nationalist ideologies, reached different parts of the world at different points in time, rolling over Latin America during the early nineteenth century and finally arriving in the Soviet Union by the end of the twentieth.

In a nutshell, the argument offered in this book proceeds along the following lines. Nationalism as a new principle of legitimacy emerged from Tilly's war-making Western states. Increasing state centralization and military mobilization led to a new contract between rulers and ruled: the exchange of political participation and public goods against taxation and the military support by the population at large. The idea of the nation as an extended family of political loyalty and shared identity provided the ideological framework that reflected and justified this new compact. It meant that elites and masses should identify with each other and that rulers and ruled should hail from the same people.

This new compact made the first nation-states of Great Britain, the United States, and France militarily and politically more powerful than dynastic kingdoms or land-based empires because they offered the population a more favorable exchange relationship with their rulers and were thus considered more legitimate. Ambitious political leaders around the world adopted this new model of statehood, hoping that they too would one day preside over similarly powerful states. These nationalists subsequently were able to establish new nation-states wherever the power configuration favored their ascent and allowed them to overthrow or gradually transform the old regime, leading to cascades of nation-state creations that altered the political face of the world over the past 200 years.

This shift from empire, dynasticism, or theocracy to national principles of legitimizing political power is a major source of war in the modern era. First, nationalists who now portrayed the ethnic hierarchies of empire as violations of the like-over-like principle resorted to arms to fight for independent nation-states. Second, newly founded nation-states competed with each other over ethnically mixed territories or over the political fate of co-nationals across the border who were ruled by ethnic others. Third, civil wars broke out when the new nation-states were captured by ethnic elites who excluded others from the political and symbolic benefits of self-rule. Such ethno-political exclusion and conflict is especially marked in states that lacked the institutional capacity and organizational bases to realize the project of nation building and to offer political participation and public goods to the population at large, rather than only to the ethnic constituencies of the dominant elites.

Nationalism thus motivated a bloody, generation-long struggle over who should rule over whom. It lasted until the like-over-like principle was realized through border changes, expulsions and ethnic cleansings, assimilation and nation building or political accommodation and power sharing between various

ethnic elites. Based on the global datasets introduced further below, we can calculate that the likelihood of war more than doubles after nationalism has gained a foothold in a political arena; and it remains high over generations after a nation-state has been founded.[1]

2 MAIN CONTRIBUTIONS

While the book tells this story of the rise and global spread of the nation-state and the waves of war it generated, it is not a history book, and it does not have a narrative structure. Rather, it explores the forces underlying these historical developments with the help of social science techniques of analysis and with large datasets that cover the entire modern world – the kind of datasets that make it possible to draw the preceding two figures. Besides introducing such new datasets, the analysis offers important substantial insights for our understanding of world history over the past two centuries. Both contributions are briefly summarized here.

2.1 Bringing power and legitimacy center stage

The book aims to show that political power and legitimacy need to move center stage in all three areas of scholarship that it addresses: on nation building and ethnic politics, on nation-state formation, and on war. It will demonstrate how particular power relations between the state and other political actors combine with their varying visions of a legitimate political order to produce different political identities, forms of statehood, and dynamics of violent conflict.

   More specifically, the book derives the political salience and legitimacy of political identities from a specific distribution of power and resources between the state and the population at large. Both ethnic group formation and nation building result from a renegotiation of the relationship between rulers and ruled during the process of political modernization (in line with Bates 1974; Wimmer 2002). Depending on how the distribution of resources and power between rulers and ruled change, political alliances form along ethnic lines, or the population at large shifts its loyalty to the state elite and identifies with the overarching national category. Ethnic groups and nations thus both represent equilibrium outcomes of the modernization process. This analysis contributes to the "constructivist" literature on ethnicity and nationalism by offering a precise, mechanism-based

---

[1] More precisely, the predicted probability of war is 1.1 percent in territories without nationalism – controlling for degrees of democratization, neighboring wars, the presence of oil resources, and political instability. This probability increases to 2.5 percent in the period after a first national(ist) organization has been founded. These figures were calculated on the basis of Model 1 in Table 4.2. Results are almost identical if we also control for levels of economic development and population size, which reduce the number of observations considerably.

analysis of the power configurations that provide either nations or specific ethnic cleavages with popular legitimacy and political meaning.

The book also introduces a power-cum-legitimacy approach to our understanding of the global spread of the nation-state. Shifts in the power relations between adherents of different ideas of legitimate statehood – dynasticism, imperial universalism, or national sovereignty – are crucial in understanding this momentous transformation of the political world over the past 200 years (in line with the general thrust of the work of Roeder 2007). The nation-state form was not universally adopted because one society after the other gradually ripened enough – as theories of modernization would have it – to finally fall as fully blossomed nations onto the garden of the inter-"national" community. Nor did the nation-state proliferate across the globe because the international system forced national sovereignty upon people after people. Similar to contagion processes, the global rise of the nation-state resulted from the concatenation of local and regional power shifts in favor of nationalists without much help from the global system. This power-configurational analysis sheds new light on a process that remains poorly understood, despite its obvious historical importance, in comparative sociology and international relations scholarship.

Finally, the book offers an analysis of war that again brings questions of political power and legitimacy to the foreground. It demonstrates that the shift of these principles of legitimacy – from empire to nation-state – is a major cause of both inter-state and civil wars over the past 200 years. This is often neglected in existing scholarship in international relations, which has paid only scarce attention to how transforming the nature of the units composing the inter-"national" system has affected war processes. The book also brings power and legitimacy to the study of civil wars that is at the core of a vast and fast-growing comparative politics literature. It demonstrates that civil wars and armed conflicts are most likely in ethnocracies that violate the principles of ethnic self-rule. Dominant political economy approaches to civil war, which focus on the conditions that make rebellion economically attractive or militarily feasible, need to be complemented with an analysis of the struggle over the power and legitimacy of the state.

2.2 New data to answer old questions

Studying nation-state formation and war has long been the exclusive domain of qualitative styles of historical research. The classic oeuvres on nationalism and the nation-state, for example, were written by historically minded social scientists such as Ernest Gellner, John Breuilly, or Michael Mann. They traced the origins of the nation-state in England, France, and the United States and then described, using examples from across the world, how it diffused over the globe. Besides these world historical narratives, entire libraries have been written on each individual trajectory of nation-state formation in the West. Others have

teased out the differences, similarities, and interlinkages between a handful of cases, often deriving big conclusions from small numbers.[2]

Most of the chapters that follow use the tools of statistical analysis to identify recurring patterns in the tapestry woven by hundreds of such specific historical threads. They will analyze newly created datasets that cover the entire world over very long periods of time and will thus allow identifying those causal mechanisms that structure more than one context and period. Such a quantitative approach based on global datasets can counterweigh against the "European provincialism" that plagues the literature on nationalism and nation-state formation, as one of its most prominent authors has trenchantly observed (Anderson 1991: xiii).[3] Emphasizing old-world developments would be less problematic if the nation-state had remained confined to the area of its origin instead of proliferating across the world, or if the earliest nation-states had indeed all been located in Europe such that those of "the rest" could be seen as belated completions of a universal sequence. However, as Anderson reminds us, the first continent to become thoroughly nationalized was the Americas, not Europe. And many non-Western nation-states came into existence before those of Europe. There is thus no reason why Holland should be given more analytical weight than Haiti, Germany more than Japan, or Belgium more than Bolivia. A quantitative approach based on global datasets gives equal weight to all cases, while allowing analysis of how they relate to each other through diffusion and imitation.

An inverse bias exists in work on ethnic politics and conflict. Here, Western scholars see themselves standing above the abyss of violence into which the leaders of many new nation-states in the East and South have thrown their populations. Studying ethnic conflicts in Africa, for example, has developed into a small research industry among comparative political scientists. But the history of Western states is punctuated by frequent episodes of ethnic cleansing and nationalist wars as well, not least during the two world wars. To see whether the West and "the rest" indeed show similar patterns of violence and war associated with the spread of nationalism and the rise of the nation-state, we thus need a perspective looking over the long run and the entire globe, rather than restricting the horizon to the world's new nation-states or the postwar period, as is the case in most comparative politics scholarship on civil war.

In order to develop such a long-term and global perspective, one needs to turn the usual relationship between data and research questions on its head. Instead of searching for new questions that have not yet been answered with

---

[2] See the well-known critique by Lieberson (1991).

[3] The articles submitted to the leading journal in the field of nationalism studies, *Nations and Nationalisms*, illustrate the disproportionate attention given to Europe: 21.5% of all manuscripts submitted since the first issue was published in 1985 were concerned with Western Europe, followed by Eastern Europe with 13.3% of the articles, then Asia, excluding the Middle East, with 12.6%, followed by Oceania with 8.7%. Only 5.4% of the articles concerned Africa, and even fewer North America (4%) or South America (2.5%).

existing datasets, new data need to be collected to answer old questions. Creating and analyzing such new datasets with global coverage represents a second major contribution that this book seeks to make to the scholarly literature. I review these data-collection efforts briefly here.

Quantitative research on civil wars often uses the readily available ethnic fractionalization index – measuring the likelihood that two randomly chosen individuals speak the same language – to see whether more diverse societies are more war-prone. Obviously, this measurement is only indirectly related to the dynamics of ethnic competition and exclusion that a long line of qualitative researchers – from John S. Furnivall (1939) to Clifford Geertz (1963), Donald Horowitz (1985), and Roger Petersen (2002) – has identified as the source of ethnic conflict. To bring quantitative research on armed conflict closer to this rich qualitative tradition of scholarship, Lars-Erik Cederman, Brian Min, and I assembled a new dataset that measures such competition and exclusion in all countries of the world and for decades of yearly observations. As Chapter 5 demonstrates, this allows us to ask more relevant questions about the nexus between ethnicity and war and to show that it is not demographic diversity that breeds violent conflict, but rather exclusionary ethno-political configurations of power.

Similarly, the relation between nation-state formation and violence cannot be properly understood with off-the-shelf datasets. These mostly take independent states as units of observation and analysis. On the one hand, this is a matter of convenience since only modern, independent states produce statistics. On the other hand, the setup of standard datasets resonates well with how both researchers and lay observers have learned to see the world – as a "family of nations" each represented by a differently colored area on a world map.

To overcome this "methodological nationalism" (Wimmer and Glick Schiller 2002), we need a universe of observations that includes colonial dependencies or pre-colonial states. Chapters 3 and 4 explore two new datasets that contain information on all territories of the world since 1816, independently of whether or not they were governed by sovereign states. This allows tracing the destiny of the world's entire population over the past two centuries and generates new insights into the dynamics of nation-state creation and its consequences for war and peace.

Another chapter reaches even deeper back into history, at the prize of focusing on two societies only. In order to see whether nation building and ethnic group formation are indeed determined by the resource and power distribution between state elites and the population at large, I have assembled data for France from the Renaissance period to the Third Republic and for the Ottoman empire from the classical age to the Young Turk revolution. These data are then fed into the formal model developed in Chapter 2. It thus takes a step beyond most other rational choice or game-theoretic models of historical processes that often rest on plausibility assumptions alone and thus are only weakly rooted in empirical data.

All five chapters, along with the long appendices that document these various data-gathering efforts, illustrate the price to pay when going beyond existing datasets. It often means struggling for each data-point, toiling through substantial amounts of sources to find that single piece of information to be filled into the cell of a spreadsheet that seems to extend its borders overnight. Are the results worth the efforts? That is for the reader to decide.

3 FOUR METHODOLOGICAL PRINCIPLES

But who would want to promote the illusion that context-free and timeless "laws of history" could ever be discovered through quantitative analysis? By adhering to the following four methodological principles, we can avoid such an overly ambitious scientism all the while identifying repeating causal dynamics in historical processes. First, we should acknowledge that causal regularity and contingency do not rule each other out, but combine to produce particular historical outcomes (King *et al.* 1994: chapter 2). It is certainly true, for example, that the assassination of the Archduke Ferdinand was a contingent event. His driver took a wrong turn into a side street of Sarajevo, where Gavrilo Princip, a pan-Serbian nationalist conspirator, happened to be on his way to lunch. He spotted and shot the archduke. This series of coincidences kindled the powder keg of World War I. But there was a powder keg waiting for a spark: a system of dyadic, uncoordinated alliances between rival states combined with the pressure of nationalist movements that sought to escape the "prisons of nations" as which they saw Eastern Europe's empires. Contemporary Europe lacks both of these conditions and it is quite unlikely that any contingent events happening on the continent will trigger a third world war at any point in the foreseeable future.

If this book seeks to explore general causal patterns, rather than historically specific chains of events, it is a matter of emphasis and choice, and not a principled stance against the role of contingency to which historical sociology has recently paid so much attention (Wagner-Pacifici 2010). While currently rather out of favor in much of sociology (*ibid.*) and comparative politics (Pierson 2003), I hope that the search for recurring long-term historical patterns can be revitalized by demonstrating that it produces robust empirical results.[4]

Second, a quantitative approach to historical processes should carefully specify the scope conditions of causal regularities in order to avoid overdrawn claims to universal validity. Some patterns may be local – they only recur throughout the history of Thailand, for example – while others are of a regional

---

[4] Development economists (Nunn 2009), neo-Malthusians working on political history (Turchin 2003) or demography (e.g. Bengtsson *et al.* 2004), and comparative political scientists studying democratization (Boix 2011) have started to explore long-term historical patterns using quantitative techniques. Some of this research has found an intellectual home in the new journal *Cliodynamics*.

scope – they exclusively shape the trajectories of former Ottoman dependencies – and still others might affect the entire world. Some causal regularities might be period specific and only effective, perhaps, after the American president Wilson had declared national sovereignty to be the right of every people on the planet. Others are valid for the entire modern age.

When searching for globally recurring causal regularities, we therefore have to pay careful attention to possible regional and period effects (Young 2009). They are best analyzed by "converting context to cause" (Collier and Mazzuca 2006) using dummy variables – investigating, for example, whether having been an Ottoman dependency is associated with a different dynamic of nation-state creation (see Chapter 3). Regional or period-specific regularities can be also discovered by sub-sample analysis, e.g. by analyzing the post-Wilsonian period in one equation and the pre-Wilsonian period in another (also in Chapter 3). To see whether the strength or even the direction of a causal relationship changes over time, key variables can be interacted with time, or we can analyze temporal sub-samples more systematically (as done by Isaac and Griffin 1989).

This book seeks to identify the causes of nation-state formation and war in the modern age, rather than those shaping particular periods and regional contexts. This is again not a matter of principle – nobody would deny that there are elements of nation-state formation and war in nineteenth-century Latin America (Centeno 2003) that are different from those of the late twentieth century Soviet Union (Beissinger 2002). Searching for regularities that hold across as many contexts as possible does come at a price, however: the story will necessarily have to be relatively abstract and general, forming a skeleton of arguments rather than a richly fleshed out and nuanced historical narrative. Whether one prefers the bones over the flesh, or whether one needs both, as lovers of mixed-method stews would argue, is largely a matter of intellectual taste, rather than of choosing between more or less "rigor," let alone empirical accuracy.

Third, this search for global patterns does not rule out that the same outcome might have multiple causes. The forces leading to ethnic conflict in Northern Ireland, to give an example, might be different from those that produced the Lebanese civil war. Such causal heterogeneity (Ragin 1989) can be discovered in a quantitative research design, for instance, with interaction effects (as in Chapter 3) or through multinomial regression analysis (see Chapter 5).

Fourth, qualitative inspection of cases and quantitative analysis of large numbers has to be combined in order to make sure that the statistical associations capture relevant mechanisms. For example, statistical analysis might discover that oil is associated with armed conflict. When investigating which cases underlie this finding, we encounter, among other "positive hits," that Mexico has oil and it has seen the Zapatista uprising in Chiapas from 1994 onwards. But the violence was not the result of a greedy hunt for oil rents. Rather, it emerged because Chiapas' entrenched *Ladino* elite had blocked land reform for generations (Collier and Lowery Quaratiello 1994). If many more such cases underlie